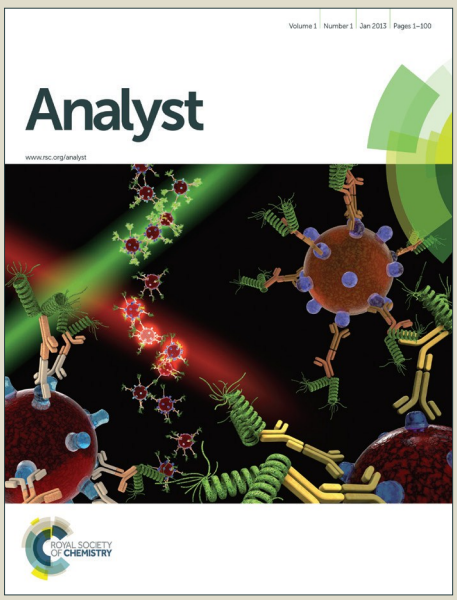
Analyst

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/analyst

Recurrence prediction in oral cancers: a serum Raman spectroscopy study

Aditi Sahu[1], Nikhila Nandakumar[1], Sharada Sawant[2], C. Murali Krishna[1]*

[1]Chilakapati lab, ACTREC, Tata Memorial Center, Kharghar, Navi Mumbai 410210 India,

[2]Vaidya lab, ACTREC, Tata Memorial Center, Kharghar, Navi Mumbai 410210 India,

* Address for correspondence

Dr. C. Murali Krishna

Scientific Officer 'F' and Principal Investigator

Chilakapati Laboratory

ACTREC, TMC

Kharghar, Sector '22'

Navi Mumbai – 410210, India.

Telephone: +91-22-2740 5039/5544

Fax: +91-22-2740 5095

Email - mchilakapati@actrec.gov.in , pittu1043@gmail.com

Keywords: Oral Cancer, Raman spectroscopy, serum, Recurrence.

Abbreviations used: SCC- Oral squamous cell carcinoma, HNSCC: Head and neck squamous cell carcinoma, RS- Raman Spectroscopy, PCA- Principal Component Analysis, PC-LDA- Principal Component Linear Discriminant Analysis, LOOCV - Leave One Out cross validation, Phe: phenylalanine

Abstract:

High mortality rates associated with oral cancers can be primarily attributed to failure of current histological procedures in predicting recurrence. Identifying recurrence related factors can lead to improved prognosis, optimized treatment and enhanced overall outcomes. Serum Raman spectroscopy has previously shown potential in diagnosis of cancers like head and neck, cervix, breast, oral cancers and also in predicting treatment response. In the present study, serum was collected from 22 oral cancer subjects [with recurrence (n=10) and no–recurrence (n=12)] before and after surgery and spectra were acquired using Raman microprobe coupled with a 40X objective. Spectral acquisition parameters were: λex = 785 nm, laser power = 30 mW, integration time: 12 s and averages: 3. Data was analyzed in patient-wise approach using unsupervised PCA and supervised PC-LDA, followed by LOOCV. PCA and PC-LDA findings suggest that recurrent and non-recurrent cases cannot be classified in before surgery serum samples; average classification efficiency of ~ 78% was obtained in after-surgery samples. Mean and difference spectra and PCA loadings indicate DNA and protein markers may be potential spectral markers for recurrence. RS of post surgery serum samples may have the potential to predict probability of recurrence in clinics, after prospective large-scale validation.

INTRODUCTION:

Head and neck cancers, which include oral cancers, are one of the leading causes of death in developing countries[1].Oral cancers are the 15th most common cancer worldwide, with an annual incidence of about 275,000 cases[2,3]. Survival of patients depends on tumor size, nodal stage, and success of initial treatment[4]. Conventional treatment for oral cancer includes surgery, radiotherapy, and chemotherapy; surgery combined with chemotherapy and radiotherapy improves overall survival. However, approximately one-third of patients treated with surgery and adjuvant therapy experience recurrence ((loco-regional, relapse, second primary and second field tumors) and/or distant metastasis. The rates of oral cancer recurrence in patients administered standard treatment vary from 18 to 76%[5], while the overall 5-year survival of 50% rate has not improved in decades.

Early detection of recurrence is clinically important[6]; patients identified at higher risk of recurrence [7,8]. Currently, the presence of cervical lymph nodes metastasis, extra capsular spread and positive histopathological margins are the important adverse prognostic factors for oral cancer[9-11] However, these existing methods are not adequate. It is known that identification of potential early markers for the development of SCC, at the level of the mucosa at risk or in serum may help in early detection of individuals at risk[12]. Tumor markers are a subset of molecules produced exclusively or in excess of normal by the pre-neoplastic or neoplastic milieu of cells. These markers may accumulate inside cells/tissues and/or released in circulation; they can aid in diagnosis, prognosis, and monitoring of treatment response. Additional markers may be secreted by recurring cancers: by neoplastic cells that remain after inadequate surgical removal, or the pre-neoplastic cells that exist as part of field cancerization[13]. Recent genetic and molecular studies on 3p14 and 9p21

chromosomal loss; p53 mutations in surgical margins have shown identification of recurrence-prone patients. A study by Reis et al has elucidated a 4-gene signature (MMP1, COL4A1, P4HA2 and THBS2) in histologically normal margins that may be predictive of oral cancer recurrence. Another study has identified 4 sub-groups of HNSCC, the subgroup with EGFR-associated profile, EMT and activation of NF-κβ signaling genes activated had poor prognosis[14-17]. Serum tumor markers may also hold promise in identifying recurrence. The tumor/recurrence-related markers with potential in determining prognosis include carcinoembryonic antigen (CEA) for colorectal cancer, CA 15-3, CEA, cMethDNA, serum testosterone levels for breast, AFP (alpha- fetoprotein) for liver, CA-125 for ovarian, prostate specific antigen (PSA) and acid phosphatase (ACP) for prostate cancer[18-24]. Several studies have also demonstrated the presence of cell-free DNA: host, tumor or viral associated as diagnostic and prognostic markers of cancers like colorectal, cervical, nasopharyngeal cancers[25-28]. Another recent study has shown utility of HPV-DNA in blood and saliva to predict recurrence in HPV-associated oral cancer patients[29]. No definite marker for recurrence prediction in non-HPV oral cancers has been established till date. Further, literature suggests that a single marker may not be efficient in detection of recurrence. A multiplex panel of several proteins and nucleic acids is being investigated for recurrence detection of several cancers[30-32]. Proteomic profiling of serum for detection of tumor/recurrence markers has been carried out for several cancers. In this context, an approach encompassing proteomics, genomics and metabolomics may be ideal for recurrence detection, and one such approach is Optical spectroscopy.

Spectroscopic studies to identify early changes using fluorescence and Raman spectroscopy (RS) have already been reported. While fluorescence spectroscopy could detect field

alterations in tumor margins, RS could detect cancer field effects (CFE)/malignancy associated changes (MAC) in oral cancer patients *in vivo*[33,34]. However their application is restricted by need for dedicated instrumentation and strict experimental conditions on-site. Serum RS has previously shown potential in detection of cancers like breast, cervical, nasopharyngeal, colorectal, head and neck and pancreatic cancers[35-40]. Our group has previously demonstrated efficacy of RS in detection of oral cancers[41,42]. Recurrence may involve reappearance of tumor- or presence of recurrence-related factors in the blood circulation. In this retrospective study, feasibility of serum RS to predict recurrence in oral cancer patients was explored. Findings are presented in the manuscript.

MATERIALS AND METHODS

Subject Details

Patients harboring primary oral squamous cell carcinoma of the oral cavity who visited the outpatient department of Tata Memorial Centre (TMC), Mumbai, India were screened for this retrospective study. A criterion of recurrence and non-recurrence was devised as follows: Subjects who reported a recurrence within 2 years of follow up were referred to as 'Recurrence subjects', while the subjects with no reported recurrence for up to 2 years of follow up were called as 'Non-recurrence subjects'. Ten subjects (n=10) fulfilled the criteria for recurrence (mean time for development of recurrence: 6 months), while n=12 fulfilled the criteria for non-recurrence. Thus, a total of 22 subjects were included in this study.

Blood was collected from these patients at 2 time points: before and after surgery. Blood samples collected after overnight fasting, prior to any surgery-related interventions was termed "before surgery" while blood collected 1 week post surgery (before any adjuvant

cancer treatment like chemoradiotherapy) was termed “after surgery”. All recruited patients were cases without prior anticancer treatment, history of malignancy, and second primary cancers. The patient's history, like age, sex, symptoms, tobacco chewing/smoking, and alcohol consumption habits, had been obtained from the hospital records and also by using a questionnaire.

Serum Separation

Five ml venous blood samples were collected from subjects with the help of a sterile injection. Samples were placed standing for 30 minutes to allow clot formation and then centrifuged at 3000 rpm for 10 minutes. The supernatant was separated and aliquoted in different tubes, and stored at -80ºC till use. One of the aliquots was utilized for Raman spectroscopic analysis while other aliquots were kept under long term storage for further/confirmatory analysis.

Raman Spectroscopy

After passive thawing, samples were subjected to Raman spectroscopy by placing 30µl volume on calcium fluoride (CaF_2) window and spectra were recorded using Fiber Optic Raman microprobe (Horiba-Jobin-Yvon, France), this system consists of laser (785 nm, Process Instruments) as an excitation source and HE 785 spectrograph (Horiba-Jobin-Yvon, France) coupled with CCD (Synapse, Horiba-Jobin-Yvon) as dispersion and detection elements, respectively. Optical filtering of unwanted noise, including Rayleigh signals, is accomplished through ‘Superhead’, the other component of the system. Optical fibers were employed to carry the incident light from the excitation source to the sample and also to collect the Raman scattered light from the sample to the detection system. Raman microprobe

was assembled by coupling a 40X microscope objective (Nikon, Japan) to the superhead . Spectral acquisition details were: excitation wavelength (λ_{ex}) = 785 nm, laser power = 30 mW. Spectra were integrated for 15 seconds and averaged over 3 accumulations. Twelve spectra were recorded from each sample.

Spectral Pre-Processing And Data Analysis

The acquired Raman spectra were corrected for CCD response and spectral contaminations from substrate and fiber signals. To remove interference of the slow moving background, first derivatives of spectra (Savitzky-Golay method and window size 3) were computed[43,44]. Spectra were interpolated in the range 700-1800 cm^{-1} since this region is an important constituent of the finger-print region. Interpolated first derivative and vector normalized spectra were then subjected to multivariate unsupervised Principal component analysis (PCA) and supervised Principal component-linear discriminant analysis (PC-LDA). In brief, Principal Component analysis (PCA) is routinely used method for data compression and visualization. It describes data variance by identifying a new set of orthogonal features, called as principal components (PCs) or factors. In LDA, the classification criterion is identified using the scatter measure of within class and between class variance. LDA can be used in conjunction with PCA (PC-LDA) to increase the efficiency of classification. The advantage of doing this is to remove or minimize noise from the data and concentrate on variables important for classification. In our analysis, significant principal components ($p<0.05$) were selected as input for LDA. In order to avoid over-fitting of the data, as a thumb rule, total number of factors selected for analysis were less than half the number of the spectra in the smallest group [45-47]. PC-LDA models were validated by Leave-one-out cross-validation (LOOCV). Leave-one-out cross validation is a type of rotation estimation, a

technique used for assessing performance of a predictive model with a hypothetical validation set when an explicit validation set is not available. Leave-one-out involves using a single observation from the original sample as the validation data, and the remaining observations as training data. This is repeated such that each observation in the sample is used once as the validation data and averaged over the rounds. Data analysis was carried out using patient-wise approach, where all spectra acquired from a single sample are averaged such that each sample is represented by a single spectrum[41]. Algorithms for these analyses were implemented in MATLAB (Mathworks Inc.) based in-house software[48] .

For spectral analysis, average spectra were computed from the background-subtracted spectra prior to derivatization for each class and were baseline-corrected by fitting a fifth order polynomial function. These baseline corrected, smoothed (Savitzky–Golay, 3) and vector-normalized spectra were the used for spectral comparisons.

Results and Discussion

The low disease-free survival rates in oral cancer patients in mainly attributed to delays in diagnosis and recurrence. Local and regional recurrence adversely influences prognosis and overall outcome of oral cancers. Current histological procedures are limited by their inability to predict recurrence. Early detection of recurrence-prone patients can lead to personalized comprehensive treatment regimens and stringent follow up, leading to a better prognosis. In cancer patients, recurrence results from i) cancer cells left behind after surgery, undetectable by histopathology (minimal residual cancer or MRC), or ii) pre-neoplastic fields which subsequently turn malignant (field cancerization or FC). The tumor, MRC or the pre-neoplastic field could secrete factors in circulation, which could be the basis for detection of

recurrence. Unlike cancers like prostate, ovary and liver, no definite serum biomarker for recurrence prediction in oral cancers is known. Serum RS has enabled detection of several cancers, including oral cancers. As recurrence in oral cancers may be associated with re-appearance of tumor and other associated factors, feasibility of recurrence prediction was explored using serum RS, before and after surgical resection of tumor in oral cancer patients.

Spectral analysis:

Mean and standard deviation spectra for Recurrence and Non-recurrence subjects before and after surgery are shown in Figure 1a-d. Major spectral features include 830 cm^{-1} and 850 cm^{-1} (Tyr doublet), 1008 cm^{-1} (Phe), 1265 cm^{-1} (Amide III), 1316 cm^{-1}, 1320 cm^{-1} and 1335 cm^{-1} (DNA related bands), 1450 cm^{-1} (CH_2 bending) and 1660 cm^{-1} (Amide I) regions. In before surgery spectra, minor differences between the recurrence and non-recurrence groups were seen at 1260 $cm^{-1,}$ 1313 cm^{-1}, 1339 cm^{-1}, 1450 cm^{-1} and 1650 cm^{-1}. These differences correspond to changes in DNA and protein in these groups. In the after surgery spectra, major differences between the recurrence and non-recurrence groups were observed at 936 cm^{-1}, 949 cm^{-1}, 1007 cm^{-1}, 1126 cm^{-1}, 1260 cm^{-1}, 1315 cm^{-1}, 1335 cm^{-1}, 1450 cm^{-1} and 1657 cm^{-1}. These differences also correspond to changes in DNA and protein across the two groups[49].

To elucidate spectral differences between the groups, difference spectra were computed by subtracting non-recurrence spectra from recurrence spectra, both before and after surgery. Positive peaks correspond to recurrence spectra while negative peaks to non-recurrence spectra. In the before surgery difference spectra (Figure 2a), prominent positive peaks were observed at 1272 cm^{-1}, 1456 cm^{-1} which indicate higher amide III and CH_2 bending of

proteins in recurrence spectra. Negative peaks were observed at 1007 cm^{-1}, 1335 cm^{-1} and 1661 cm^{-1}, which correspond to a lower phenylalanine, DNA and amide I content in the recurrence spectra. Our previous studies have shown that higher protein and DNA features are observed in tumor sera spectra, with respect to healthy controls. Thus, these biochemical features may correspond to tumor-related factors. As these factors predominate in both groups, detection of recurrence-related factors, if any, may not be easy in before surgery spectra.

In the after surgery difference spectra (Figure 2b), positive peaks are observed at 1009 cm^{-1} (Phe), 1255 cm^{-1}, 1280 cm^{-1} (amide III), 1342 cm^{-1} (DNA), 1450 cm^{-1} (CH2 bending) and 1677 cm^{-1} (amide I), indicating an overall high Phe, DNA and protein content in the recurrence spectra. As these samples were collected after surgical resection of tumor, both groups now contain mainly normal serum constituents. The additional DNA and protein signals could originate from either minimal residual cancer or field cancerization. This corroborates with findings that demonstrate high circulating DNA levels in the recurrence group, even higher than the primary cancer group; along with up-regulation of several proteins in the sera of recurrence patients. Thus, this additional DNA and protein content could be ascribed to recurrence-related factors. Further, a proteomic study has delineated tumor and recurrence-related factors. Different protein peak patterns were observed after MALDI-TOF-MS of recurrent and primary ovarian cancers. Thus, tumor and recurrence-related factors may have differential origin and basis[50].

Multivariate analysis

Spectral features indicate differences between the recurrence and non-recurrence groups. To explore the feasibility of classifying these groups, multivariate analysis using unsupervised principal component analysis (PCA) and supervised principal component-based linear discriminant analysis (PC-LDA) was carried out. PC-LDA results were further validated by Leave-one-out cross validation (LOOCV). Patient-wise approach (all spectra from a sample averaged to yield a representative spectrum) was adopted for data analysis (analyst ref). First, differences between the recurrence and non-recurrence groups were analyzed in *before surgery* sera. Next, the same approach was adopted for *after surgery* sera. Results are presented in the form of scatter plots (PCA, PC-LDA) and confusion matrix (PC-LDA, LOOCV).

Investigating differences in serum *before surgery*

In the first step, 23 spectra from 10 recurrence and 11 non-recurrence subjects were subjected to PCA. Scores of factor 2 and 3 were explored for classification. The loadings of factor 2 and 3, and the scatter plot are shown in Figure 3. The scatter plot indicates large overlap between the recurrence and non-recurrence groups. In the PCA loadings of *before surgery* spectra, factor loading 2 has peaks at 948, 1010, 1337, 1450 and 1660 cm^{-1} , thus features of Phe, and mainly proteins (CH_2 bending, amide III and amide I region) are contributed by factor 2. Factor loading 3 has peaks at 736, 934, 1110, 1156, 1354, 1398, 1502, 1522, 1645, 1743 cm^{-1} which can be broadly attributed to contributions from amide I and ester regions. The subtle differences in the before surgery PCA can therefore be ascribed mainly to protein content.

As PCA is not a classification tool but is used for data compression and visualization to indicate trends in the data, PC-LDA was employed to explore classification between the groups. Three factors were used for the analysis, accounting for ~81% correct classifications. Scores of factor 2 and 3 were employed to obtain scatter plots, as shown in Figure 4. As seen in PCA, overlap between the two groups was observed.

As seen in PC-LDA confusion matrix (Table 1a), 9/10 recurrence spectra were correctly classified, while 8/11 non-recurrence spectra were correctly classified. As PC-LDA is a supervised approach, leave-one-out-cross-validation (LOOCV) was carried out to evaluate the results obtained by PC-LDA. On LOOCV (Table 1b), 7/10 recurrence spectra and 4/11 non-recurrence spectra were correctly classified, to yield a classification efficiency of 70% and 36%, respectively. A large number of misclassifications of non-recurrence group with the recurrence group were observed. However, a minor tendency of classification was observed for recurrence group.

Mean spectra analysis indicates high DNA and protein features (tumor-related factors) in both groups. There is no classification between the two groups, as seen in PCA and PC-LDA results. Before surgery, both recurrence and non-recurrence patients' sera comprise of a) normal serum constituents and b) tumor related factors. The recurrence group may also contain some recurrence related factors, arising due to putative presence of pre-neoplastic fields. However, due to abundance of normal and tumor factors, detection of any recurrence-related factors may be difficult in before surgery samples. Therefore, after surgery samples were also analyzed.

Investigating differences in serum *after surgery*

Recurrence-related differences could not be detected in before surgery samples. This may be possibly attributed to the presence of additional tumor-associated factors (along with normal serum constituents and recurrence-related factors, if any). These factors may be eliminated from circulation by surgical excision of tumor, and may facilitate detection of recurrence-related factors, if any. Further, as formerly stated, recurrence can develop due to two main reasons, minimal residual cancer (MRC) and field cancerization (FC). Thus, recurrence prediction is based on factors arising from MRC, FC or both. As MRC factors can only be detected post-surgical excision of tumor, blood samples collected post-surgery were also analyzed for recurrence detection.

Serum collected post-surgery was also subjected to Raman spectroscopy. Thus, 22 spectra from 10 recurrent subjects and 12 non-recurrent subjects were subjected to PCA. Scores of factor 2 and factor 3 were employed to explore classification. The loadings of factor 2 and 3 and the scatter plot are shown in Figure 5. Scatter plot indicates two almost distinct groups, corresponding to recurrence and non-recurrence sera. In the PCA loadings after surgery, factor loading 2 shows bands at 754, 785, 920, 1008, 1118, 1198, 1340, 1450, 1663 cm^{-1} attributed to Phe, DNA bases, CH_2 bending and amide I of proteins while factor loading 3 shows bands at 1014, 1342, 1451, 1624 cm^{-1} indicating predominance of components like Phe, DNA bases, and protein (CH_2 bending and amide). Thus, the differences between the PCA of after surgery serum samples can be attributed to differential DNA and protein content in the groups.

PC-LDA was carried out with 3 factors that accounted for ~ 82 % classifications. The scatter plot for PC-LDA (score of factor 1 vs. score of factor 2) is shown in Figure 6. As in PCA, 2 well-separated groups were observed. The confusion matrix in Table 2a for PC-LDA yields

9/10 correct predictions for recurrence group while 9 /12 correct classifications for non-recurrence group. LOOCV results, presented in Table 2b indicates 8/10 and 8/12 correct predictions for recurrence and non-recurrence, to yield a classification efficiency of 80% and 75%, respectively. Thus, recurrence and non-recurrence groups could be classified with average classification efficiency of ~78%.

After removal of tumor, normal serum constituents could be major contributors to Raman spectra. However, some recurrence related factors persisting in sera of recurrence subjects may enable classification between the two groups. Further, as influence of confounding tumor-related factors may have been removed by surgical excision of tumor, these recurrence-related factors may have played a major role in classification of the groups. Thus, blood from oral cancer patients after surgery may have the potential to identify those at high risk of developing recurrence. Recurrence may have been detected by factors arising from MRC, FC or both[51,52].

A study on cytokeratin profiling has shown elevated levels of Tissue polypeptide antigen (TPA) assay which detects fragments of cytokeratin 8, 18 and 19 in the recurrence group, both before and after surgery. A similar aberrant cytokeratin expression was also encountered in the tumor-associated normal mucosa of the same patients, indicating that presence of cytokeratin post surgery could be attributed to these factors in circulation[13]. Additionally, the differences could be attributed to recurrence-related factors like elevated protein and DNA seen in the difference spectra and PCA factor loadings. This is also corroborated by several molecular studies[25-32].

Although the exact underlying reason for differences between the recurrence and non-recurrence groups can only be speculated in the current study, the finding that the recurrence-prone oral cancer patients can be identified using serum RS has important clinical implications. Conventionally, oral cancer patients with poor prognosis or lymph node metastasis are administered adjuvant chemo-radiotherapy to ensure a comprehensive treatment and prevent recurrence. If recurrence-prone subjects could be identified by the current methodology 1 week post surgical excision of tumor, similar comprehensive treatment regimens can be planned for such patients. Irrespective of histopathological grading and/or nodal metastasis (which may prove to be inadequate), adjuvant treatment post surgery can be administered to such patients. Stringent bi-monthly or monthly follow-ups along with regular imaging modalities to detect even any occult suspicious lesions can be planned. In event of lesion confirmation, treatment options can then be weighed to promote patient life-quality and decrease morbidity.

CONCLUSIONS

Low disease-free survival rates in oral cancer patients is attributed to recurrence. Current histological procedures are limited by their inability to predict recurrence. Early detection of recurrence-related factors can lead to less morbidity, increased disease free survival and better quality of life for patients. In this retrospective study, feasibility of differentiating serum of recurrence and non- recurrence oral cancer patients using RS was explored. Serum (previously collected) was analyzed for 2 time points: before surgery (prior to any anti-cancer treatment) and 1 week after surgical excision of the tumor (prior to any adjunctive chemotherapy or radiotherapy). Prominent changes with respect to DNA and proteins in the mean and difference spectra and loadings indicate that these molecules are major contributors

to recurrence spectra. These findings corroborate with the existing literature that suggests increased cell-free DNA: host, tumor or viral associated as diagnostic and prognostic markers and up-regulation of several proteins in circulation in cancers like colorectal, cervical, nasopharyngeal cancers. Thus recurrence may indeed be associated with higher secretion of DNA and proteins by the remaining cancer i.e. minimal residual cancer or the pre-neoplastic field existing in the cancer patient. Multivariate analysis indicates recurrence group can be identified after surgery, PC-LDA followed by LOOCV distinguished recurrence and non-recurrence groups with an efficiency of ~ 77%. The observed differences between recurrence and non-recurrence groups seen more prominently in after surgery spectra could be due to removal of confounding tumor-related factors. The current exploratory study highlights the feasibility of identifying recurrence-prone patients. A large-scale validation study with a huge sample size can help in establishing Raman spectral markers for recurrence, which could further be confirmed by biological assays, prospectively leading to implementation of this method in clinics. Although identification of patients at high-risk for recurrence using serum RS cannot predict localization of recurrent tumor, it can serve as a preliminary test. On the basis of these results, regular Imaging (modalities like PET, CT or MRI) followed by more comprehensive adjuvant treatment decisions and stringent follow-ups may improve overall outcome of the disease.

ACKNOWLEDGEMENT:

The Raman spectrometer employed in the study was procured from DBT project BT/ PRI11282/MED/32/83/2008 (Development of in vivo laser Raman spectroscopy methods for diagnosis of oral precancerous and cancerous conditions), Department of Biotechnology,

Government of India. The authors would also like to acknowledge Mr. Shridhar Nadkar for his help in sample collection and storage.

REFERENCES:

1. Z. Khan, J. Tönnies, and S. Müller, *J. Can Epidem* 2014, 2014 doi:10.1155/2014/394696

2. J. Ferlay, I. Soerjomataram, M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman and F. Bray, *GLOBOCAN 2012* version 1.0, Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 11 Lyon, France: International Agency for Research on Cancer; 2013.

3. A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, *CA Cancer J. Clin*, 2011, 61, 69–90.

4. P. Garzino-Demo, A. Dell'Acqua, P. Dalmasso, M. Fasolis, G. M. La Terra Maggiore, G. Ramieri, S. Berrone, M. Rampino, M. Schena, *J. Craniomaxillofac Surg.* 2006, 34, 344–350, 10.1016/j.jcms.2006.04.004

5. S. D. Da Silva, M. Hier, A. Mlynarek et al. *Front Pharmacol.* 2012;3:149.

6. L. Kostakoglu, R. Fardanesh and M. Posner, *Oncologist*, 2013;18(10):1108–1117.

7. M.Partridge, S-R Li, S. Pateromichelakis, R. Francis, E. Phillips, X. H. Huang, F. Tesfa-Selase, and J. D. Langdon, *Clin Cancer Res*., 2000 6:2718-2725.

8. G. R. Thomas, H. Nadiminti and J. Regalado, *Int J Exp Pathol.*, 2005, 86: 347–363.

9. J. A. Woolgar, S. Rogers, C. R. West, R. D. Errington, J. S. Brown, and E. D. Vaughan, *Oral Oncol*., 1999, 35: 257–265.

10. J. S. Greenberg, A. K. El-Naggar, V. Mo, D. Roberts and J. N. Myers, *Cancer* 2003, 98, 508–515, 10.1002/cncr.11526

11. C. R. Leemans, R. Tiwari, J. J. Nauta, I. van der Waal and G. B. Snow, *Cancer*, 1994, 73, 187–190, 10.1002/1097-0142

12. I. J. Dhooge, M. De Vos and P. B. Van Cauwenberge, *The Laryngoscope* 1998, 108: 250–256

13. S. S. Sawant, D. A. Chaukar, S. S. Joshi, P. P. Dange, S. Kannan, S. Kane, A. K. D'Cruz, M.M. Vaidya, *Oral Oncol 2011* , 47(2): 114 - 120

14. M. P. Rosin, X. Cheng, C. Poh, W. L. Lam, Y. Huang and J. Lovas, *Clin Cancer Res.* 2000;6:357–62.

15. A. P. Graveland, P. J. Golusinski, M. Buijze, R. Douma, N. Sons, D. J. Kuik, E. Bloemena, C. R. Leemans, R. H. Brakenhoff and B. J.M. Braakhuis, *Int. J. Cancer*, 2011, 128: 1852–1859. doi: 10.1002/ijc.25523

16. P. P. Reis, L. Waldron, B. Perez-Ordonez, M. Pintilie, N. N. Galloni, Y. Xuan, N. K. Cervigne, G. C. Warner, A. A. Makitie, C. Simpson, D. Goldstein, D. Brown, R. Gilbert, P. Gullane, J. Irish, I. Jurisica, S. Kamel-Reid, *BMC Cancer*, 2011;11:437. doi: 10.1186/1471-2407-11-437

17. C. H. Chung, J. S. Parker, G. Karaca, J. Wu, W. K. Funkhouser, D. Moore, D. Butterfoss, D. Xiang, A. Zanation, X. Yin,W. W. Shockley, M. C. Weissler, L. G. Dressler, C. G. Shores, W. G. Yarbrough, C. M. Perou, *Cancer Cell*, 2004, 5, 489-500.

18. F. Lumachi, S. M. M. Basso, U. Basso, Breast Cancer Recurrence: Role of Serum Tumor Markers CEA and CA 15-3, Methods of Cancer Diagnosis, Therapy and Prognosis, 2008,1,109-115.

19. M. Han, S. Piantadosi, M. L. Zahurak, L. J. Sokoll, D. W. Chan, J. I. Epstein, P. C. Walsh, A. W. Partin, *Urology*, 2001;57:707–711. doi: 10.1016/S0090-4295(00)01073-6.

20. T. J. Wallace, T. Torre, M. Grob, J. Yu, I. Avital, B. L. D. M. Brücher, A. Stojadinovic, Y. G. Man, *J Cancer*, 2014;5(1):3–24.

21. A. W. Partin, , G. E. Hanks, E. A. Klein, J. W. Moul, W. G. Nelson, and H. I. Scher, *Oncology*, 2002

22. E. Tan, N. Gouvas, R. John Nicholls, P. Ziprin, E. Xynos, P. P. Tekkis, *Surgical Oncology*, 2009, 18:1, 15-24, ISSN 0960-7404, http://dx.doi.org/10.1016/j.suronc.2008.05.008.

23. Q. Lai, F. Melandro, R. S. Pinheiro, et al., *Int J Hepatol*, 2012, 2012, doi:10.1155/2012/893103

24. C. M.Sturgeon, E. P. Diamandis, Use of Tumor Markers in Testicular, Prostate, Colorectal, Breast, and Ovarian Cancers, The National Academy of Clinical Biochemistry Laboratory Medicine Practice Guidelines, 2009, Washington DC.

25. A. Widschwendter, A. Blassnig, A. Wiedemair, E. Muller-Holzner, H. M. Muller, C. Marth, *Cancer Lett.* 2003, 202(2):231–9

26. T. B. Hao, W. Shi, X. J. Shen, J. Qi, X. H. Wu, Y. Wu, Y. Y. Tang and S Q Ju, *Br J Cancer* 2014, 111, 1482-1489

27. T.Behne and M. S.Copur, *Int J Hepatol* 2012, 859076.

28. K. Shotelersuk, C. Khorprasert, S. Sakdikul, W. Pornthanakasem, N. Voravud, A. Mutirangura, *Clin Cancer Res*. 2000, 6:1046–1051.

29. S. M. Ahn, J. K. Chan, Z. Zhang, et al. *JAMA Otolaryngol Head Neck Surg.* 2014;140(9):846-854. doi:10.1001/jamaoto.2014.1338.

30. L. Chen., M. Chang, Y. Dai, K.X. Chai, L.Dyrskjøt, M. Sanchez-Carbayo, T. Szarvas, E. C. Zwarthoff, V. Lokeshwar, C. Jeronimo, A. S. Parker, S.Ross, M. Borre, T.F. Ørntoft, T. Jaeger, W. Beukers, L. E. Lopez, R. Henrique, P. R. Young, V. Urquidi, S. Goodison, and C. J. Rosser *Cancer Epidemiol Biomarkers Prev* 2014, 23:1804-1812.

31. Q. Long, B. A. Johnson, A. O. Osunkoya, Y. H. Lai, W. Zhou, M. Abramovitz, M. Xia, M. B. Bouzyk, R. K. Nam, L. Sugar, A. Stanimirovic, D. J. Williams, B. R. Leyland-Jones, A. K. Seth, J. A. Petros, C. S. Moreno, *Am J Pathol.* 2011, 7(1):46–54.

32. M. Chatterjee,G. Dyson, N. K. Levin,J. P. Shah, R.Morris, A. Munkarah, M. A. Tainsky, *Cancer Biomark 2012, 11 (23), 59-73.*

33. C. F. Poh, L. Zhang, D. W. Anderson, J. S. Durham, P. M. Williams, R. W. Priddy, K. W. Berean, S. Ng, O. L. Tseng, C. MacAulay, M. P. Rosin, *Clin Cancer Res.* 2006, 12(22):6716-22.

34. S. P. Singh, A. Sahu, A. Deshmukh, P. Chaturvedi and C. M. Krishna, *Analyst*, 2013, DOI: 10.1039/c3an36761d

35. A. T. Harris, A. Lungari, C. J. Needham, S. L. Smith, M. A. Lones, S. E. Fisher, X. B. Yang, N. Cooper, J. Kirkham, D. A. Smith, D. P. Martin-Hirsch and A. S. High, *Head Neck Oncol.,* 2009, 1:34, DOI:10.1186/1758-3284-2-26.

36. J. L. Pichardo-Molina, C. Frausto-Reyes, O. Barbosa-Garcia, R. Huerta-Franco, J. L. Gonzalez-Trujillo, C. A. Ramirez-Alvarado, G. Gutierrez-Juarez and C. Medina-Gutierrez, *Lasers Med. Sci*., 2007,22,229–236.

37. S. Feng, R. Chen, J. Lin, J. Pan, G. Chen, Y. Li, M. Cheng, Z. Huang, J. Chen and H. Zeng, *Biosens Bioelectron*., 2010, 25(11), 2414–2419.

38. X. Li, T. Yang and S. Li, *Appl Opt.* 2012, 51(21),5038-43.

39. J. L. González-Solís, J. C. Martínez-Espinosa, L. A. Torres-González, A. Aguilar-Lemarroy, L. F. Jave-Suárez, P. Palomares-Anda, *Lasers Med Sci* 2014, 29, 3, 979-985.

40. G. Wang, R. J. Lipert, M. Jain, et al., *Analytical Chemistry*. 2011;83(7):2554–2561.

41. A. Sahu, S. Sawant, H. Mamgain and C. Murali Krishna, *Analyst* 138, 4161-4174, 2013.

42. A. Sahu, S. Sawant, S. Talathi-Desai, C. Murali Krishna, *Proc of SPIE 2014*, 89390E doi: 10.1117/12.2034941.

43. A. Nijssen, K. Maquelin, P. J. Caspers, T. C. B. Schut, M. H. A. Neumann and G. J. Puppels, *J. Biomed. Opt.* 2007,12, 034004-1 – 034004-7

44. S. Koljenovic, L. P. Choo-Smith, T. C B. Schut, J. M. Kros, H. J. Berge and G.J. Puppels, *Lab Invest., 2002, 82*, 1265-1277.

45. P. Crow, B. Barrass, C. Kendall, M. Hart-Prieto, M. Wright, R. Persad and N. Stone, *Br J Cancer 2005, 92,* 2166 –2170.

46. T. C. B. Schut, R. Wolthuis and G. J. Caspers, *J. Raman Spectrosc.*,2002, *33(7)*, 580-585.

47. T. J. Harvey, E. Gazi, A. Henderson, R. D. Snook, N. W. Clarke, M. Brown and P. Gardner, *Analyst*, 2009, *134*, 1083-1091.

48. A. D. Ghanate, S. Kothiwale, S. P. Singh, D. Bertrand and C. M. Krishna, *J Biomed Opt*, 2011, 16, 025003.

49. F. S. Parker, *Applications of Infrared, Raman and Resonance Raman spectroscopy in Biochemistry*, Plenum Press, Newyork and London, 1983.

50. Y. Wang, J. J. Yu, T. Zhu, L. Xu, M. Xu, Y. Z. Huang, H. Pu, C. Q. Yu, *Asian Pac J Trop Med.* 2012;5(7):573-6.

51. B. J. Braakhuis, R. H. Brakenhoff, C. R. Leemans, *Ann Oncol. 2012*;23(Suppl 10):173–177.

52. M. G. C. T. V. Oijen and P. J. Slootweg,Cancer Epidemiol., *Biomarkers Prev.*, 2000, 9, 249–256.

Table 1. PC-LDA analysis for *before surgery* recurrence and non-recurrence samples.

Confusion matrix for a) PC-LDA and b) Leave-one-out cross validation (LOOCV)

	BS- R	BS- NR	total	% efficiency
BS- R	9	1	10	90
BS- NR	3	8	11	72.7

(a)

	BS- R	BS- NR	total	% efficiency
BS- R	7	3	10	70
BS- NR	7	4	11	36.3

(b)

Table 2. PC-LDA analysis for *after surgery* recurrence and non-recurrence samples. Confusion matrix for a) PC-LDA and b) Leave-one-out cross validation (LOOCV)

	AS- R	AS- NR	total	% efficiency
AS- R	9	1	10	90
AS- NR	3	9	12	75

(a)

	AS- R	AS- NR	total	% efficiency
AS- R	8	2	10	80
AS- NR	4	8	12	75

(b)

Figure Legends:

Figure 1. Mean and standard deviation spectra of before surgery samples a) recurrence b) non-recurrence and after surgery samples c) recurrence d) non-recurrence

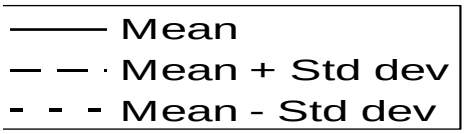


Figure 2. Difference spectra of recurrence-non-recurrence serum

a) Before surgery, b) After surgery

Figure 3. PCA for *before surgery* serum samples

a) Loadings of factor 2, b) Loadings of factor 3, c) Scatter plot.

Figure 4. PC-LDA for *before* surgery serum samples

a) Scree plot b) Scatter plot

Figure 5. PCA for *after surgery* serum samples

b) Loadings of factor 2, b) Loadings of factor 3, c) Scatter plot.

Figure 6. PC-LDA for *after* surgery serum samples

b) Scree plot b) Scatter plot

Figures

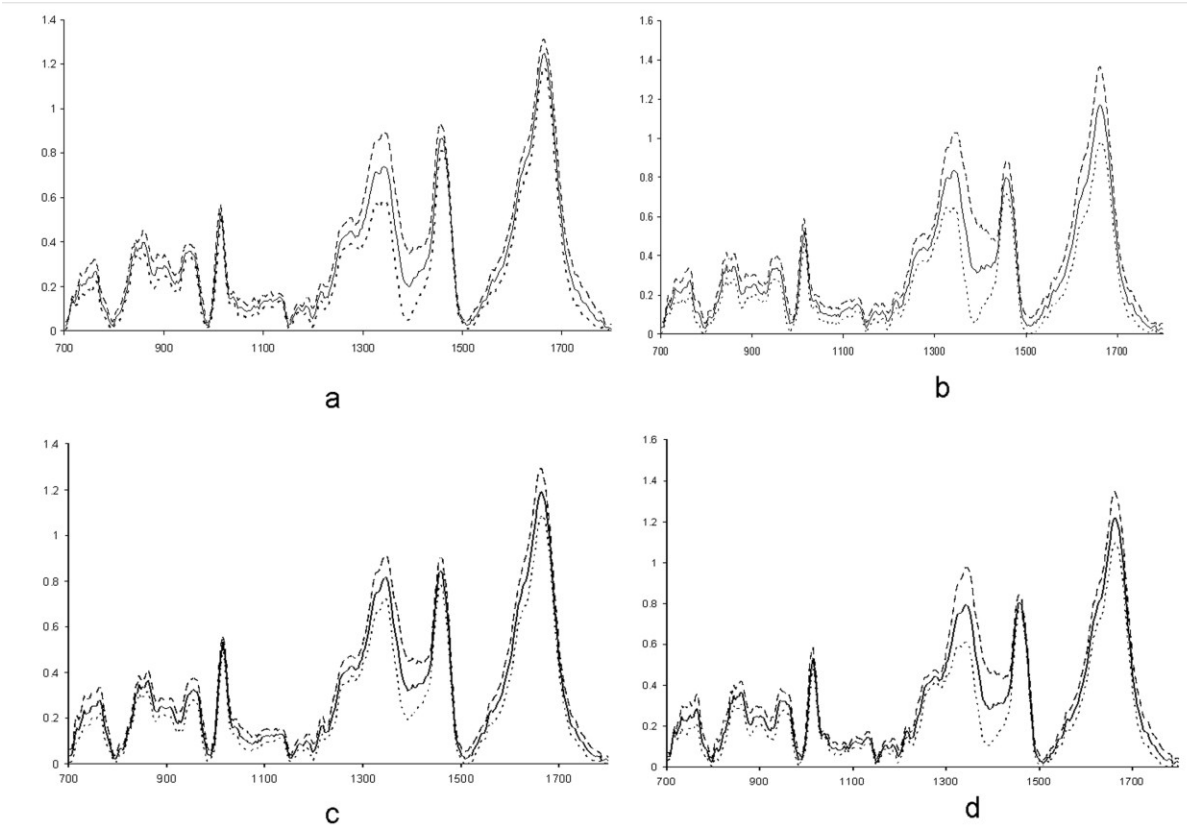


Figure 1. Mean and standard deviation spectra of before surgery samples a) recurrence b) non-recurrence and after surgery samples c) recurrence d) non-recurrence

(——— Mean, --- --- --- Mean + Standard deviation, - - - - - - Mean - Standard deviation)

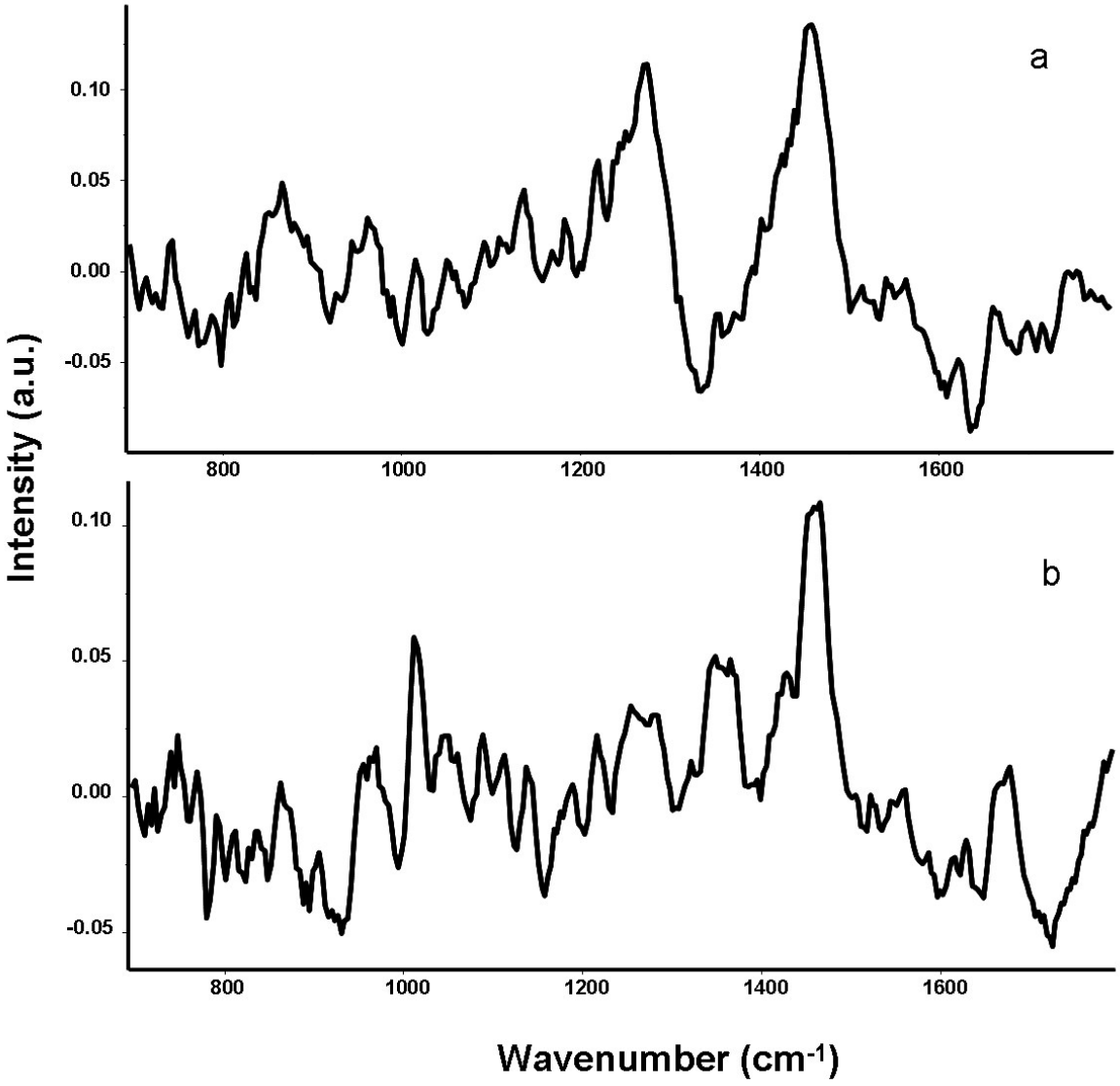


Figure 2. Difference spectra of recurrence-non-recurrence serum

b) Before surgery, b) After surgery

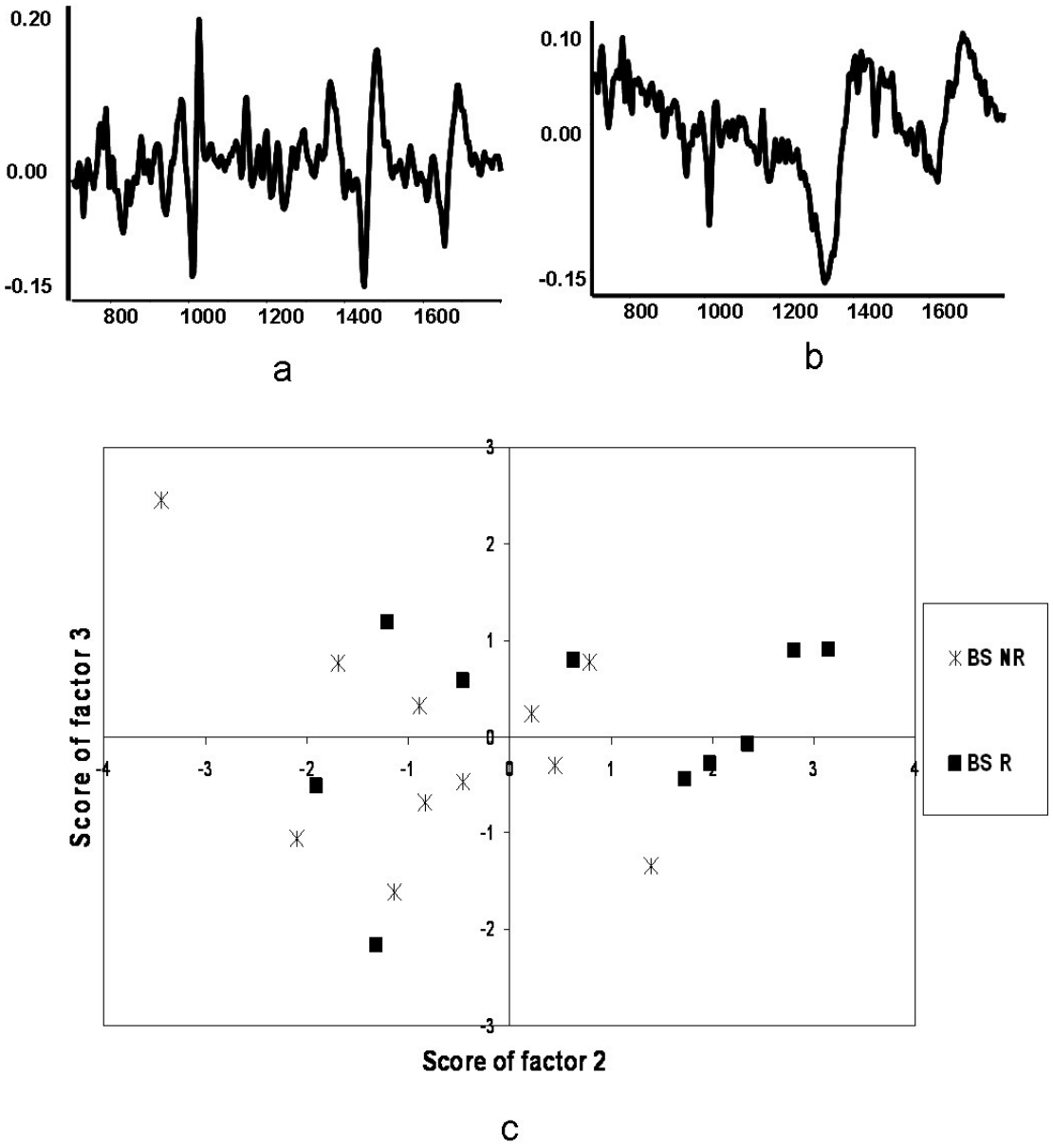
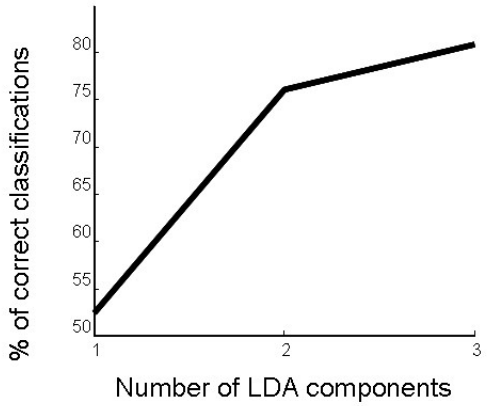
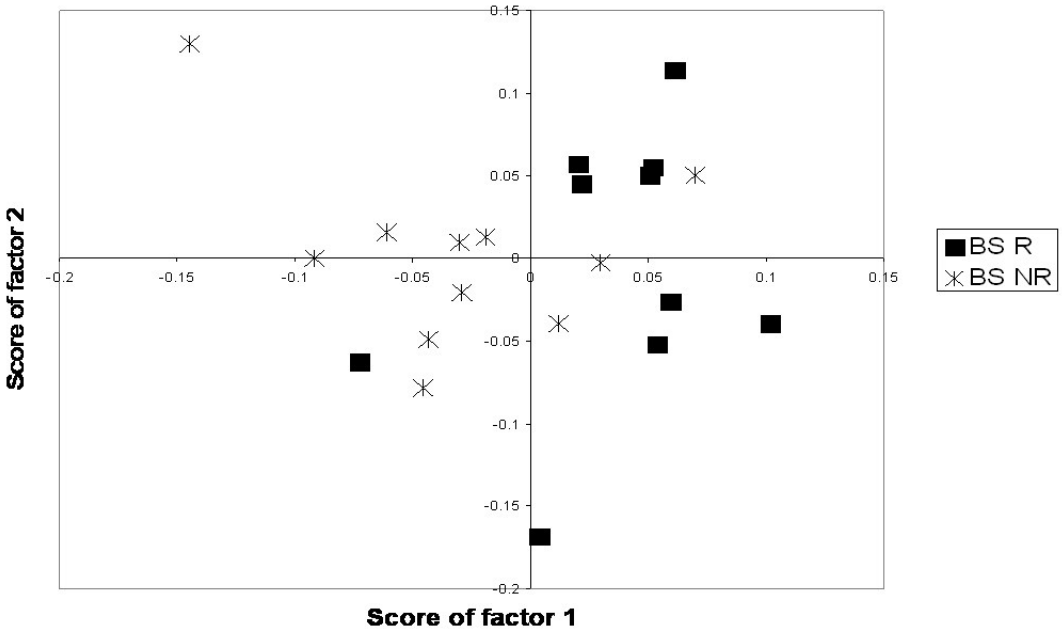


Figure 3. PCA for *before surgery* serum samples

c) Loadings of factor 2, b) Loadings of factor 3, c) Scatter plot.



a



b

Figure 4. PC-LDA for *before* surgery serum samples

c) Scree plot b) Scatter plot

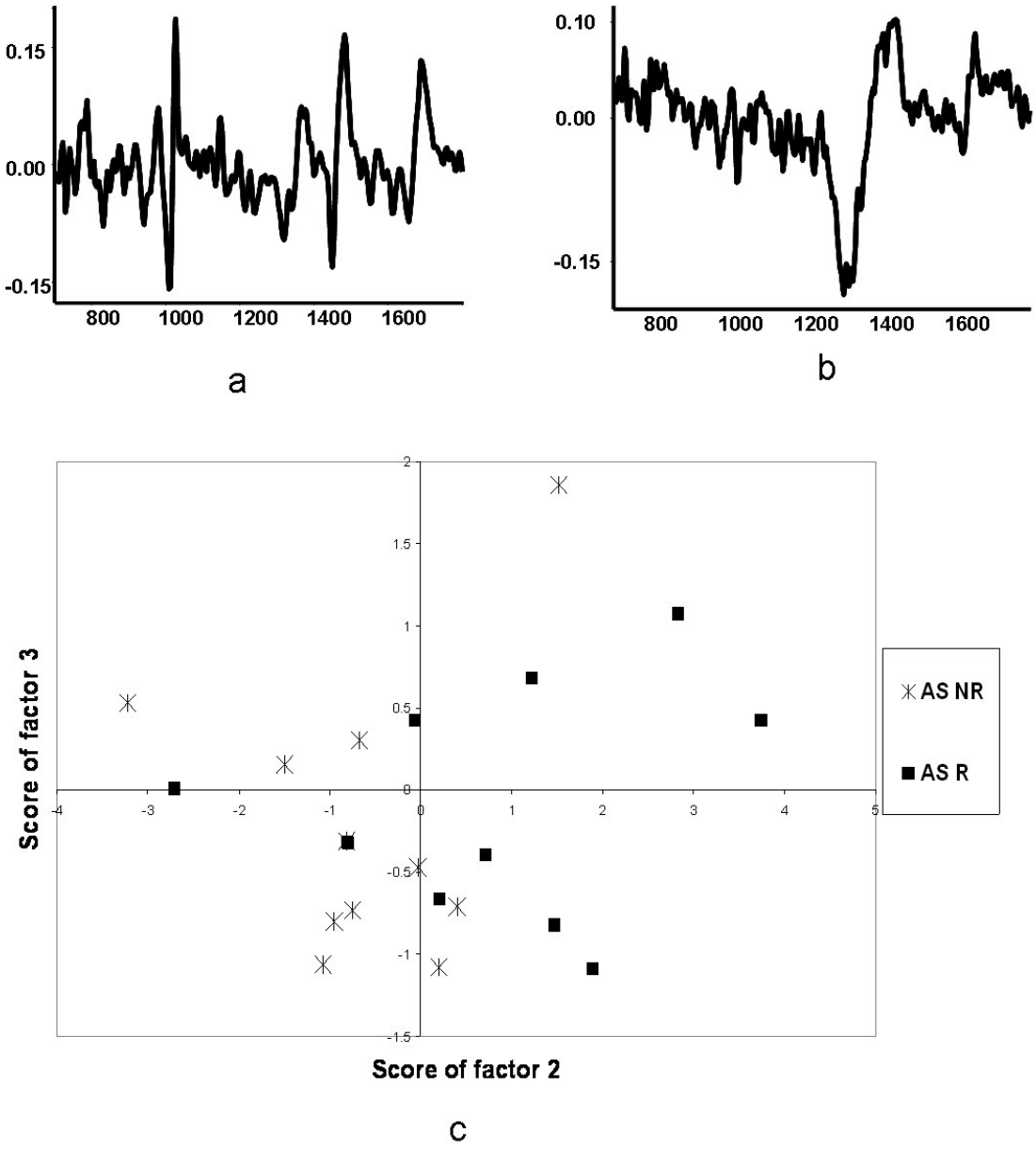


Figure 5. PCA for *after surgery* serum samples

d) Loadings of factor 2, b) Loadings of factor 3, c) Scatter plot.

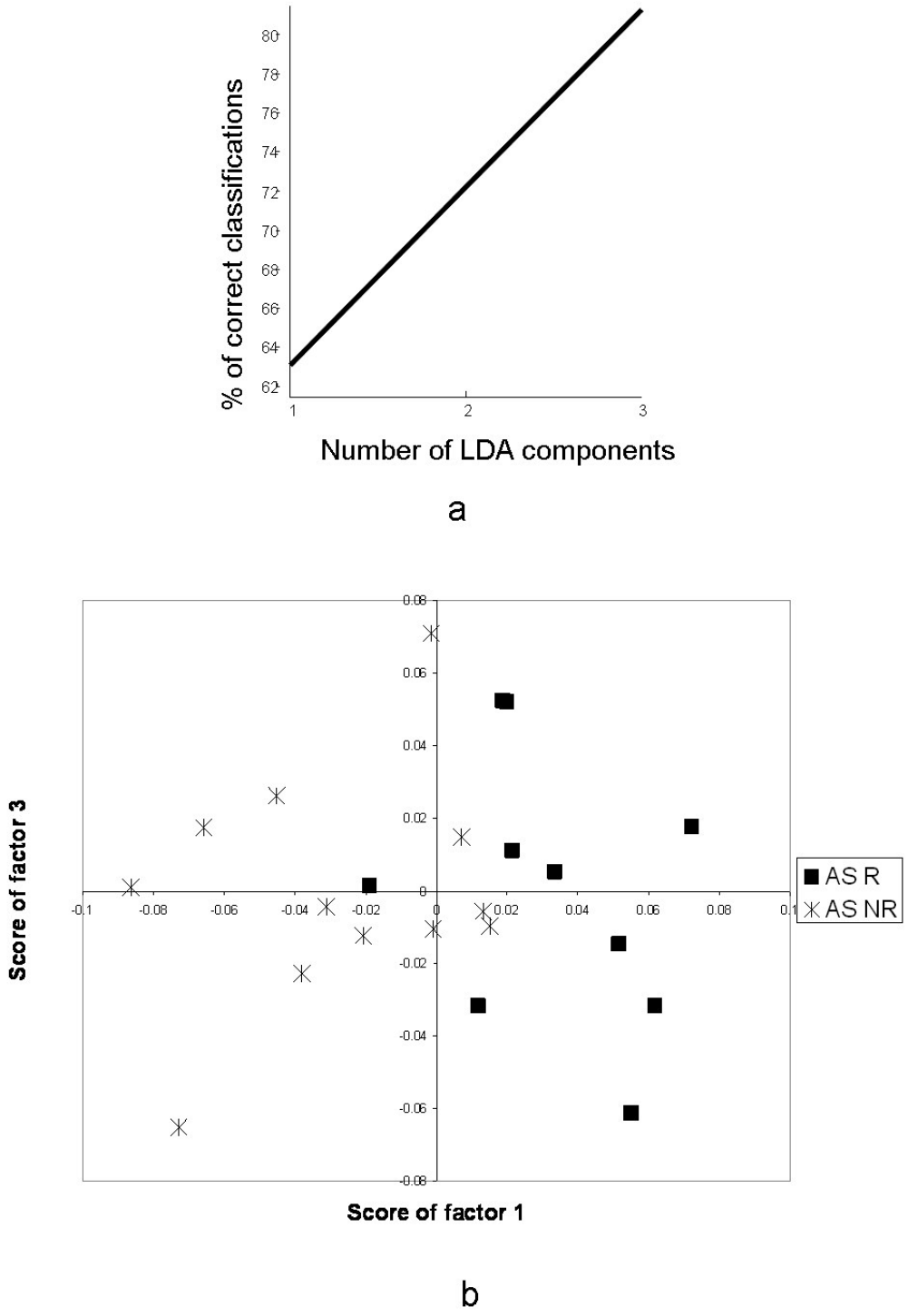


Figure 6. PC-LDA for *after* surgery serum samples

d) Scree plot b) Scatter plot